Analyse 7 - Fonctions de deux variables

1 Définition, domaine de définition et graphe

► Intuitivement, une fonction est une quantité qui varie en fonction d'un paramètre. Il se peut que la quantité étudiée dépende de plusieurs paramètres, on parle alors de fonction de plusieurs variables. Nous allons étudier dans ce chapitre les fonctions de deux variables.

Exemples :

- Fonctions coordonnées : $(x, y) \mapsto x$ et $(x, y) \mapsto y$.
- Distance entre x et y : $(x, y) \mapsto |x - y|$.
- Altitude d'un point d'une carte topographique : $(x, y) \mapsto \text{alt}(x, y)$.

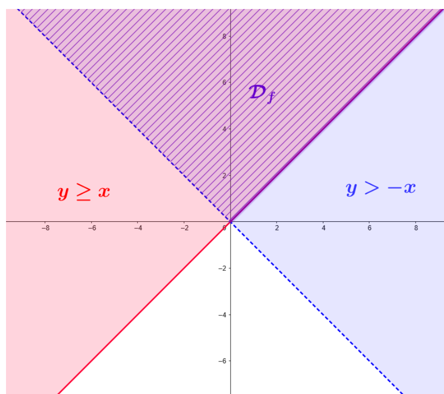
Remarque. *Comme pour les fonctions à une variable, les fonctions de deux variables n'ont qu'un seul espace de départ :* $\mathbb{R}^2$ *(ou une sous-partie de* $\mathbb{R}^2$*).*

► Le **domaine de définition** d'une fonction de deux variables est une partie de $\mathbb{R}^2$, c'est l'ensemble des points $(x, y) \in \mathbb{R}^2$ tels que $f(x, y)$ est bien défini.

Exemples :

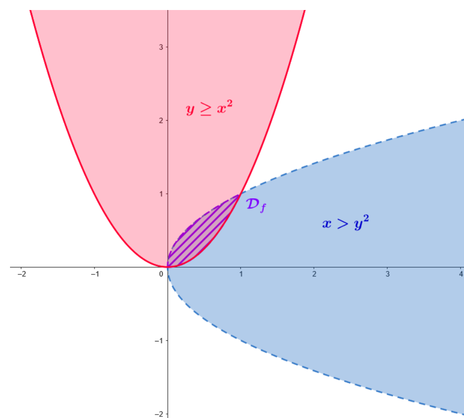
- $f : (x, y) \mapsto ye^x$, on a $\mathcal{D}_f = \mathbb{R}^2$.
- $f : (x, y) \mapsto x \ln(x + y) - y\sqrt{y - x}$.

$$\mathcal{D}_f = \left\{(x, y) \in \mathbb{R}^2 \mid y > -x \quad \text{et} \quad y \geqslant x\right\}$$



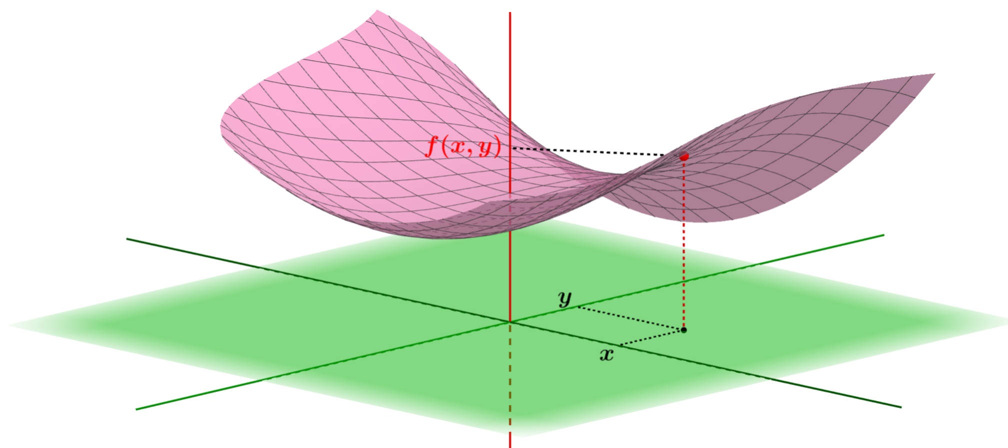
- $f : (x, y) \mapsto \ln(x - y^2) - 2\sqrt{y - x^2}$.

$$\mathcal{D}_f = \left\{(x, y) \in \mathbb{R}^2 \mid y \geqslant x^2 \quad \text{et} \quad x > y^2\right\}$$



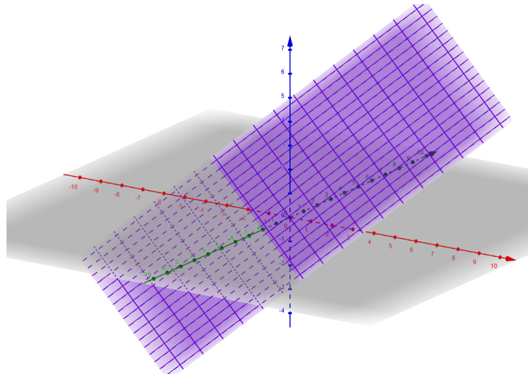
► On représente les fonctions d'une variable graphiquement : la courbe représentative d'une fonction est une partie de $\mathbb{R}^2$, c'est l'ensemble des points de coordonnées $(x, f(x))$ avec x qui parcourt l'ensemble de définition de la fonction.
De la même manière, on peut représenter graphiquement les fonctions de deux variables : le **graphe** d'une telle fonction est une partie de $\mathbb{R}^3$ (surface) :

$$\boxed{\boldsymbol{\Gamma_f = \{(x, y, f(x, y)) \mid (x, y) \in \mathcal{D}_f\}}}.$$



Exemples :

- On considère la fonction $f : (x, y) \mapsto x + y + 1$. Γ_f est le plan d'équation $z = x + y + 1$.

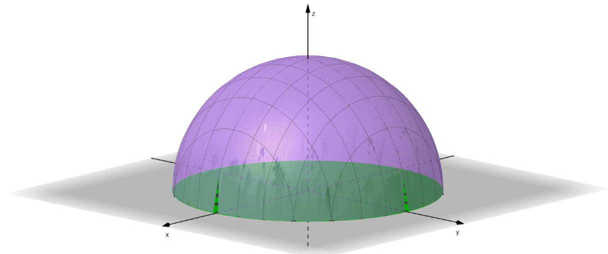


- On considère la fonction $f : (x, y) \mapsto \sqrt{1 - x^2 - y^2}$.
  f est définie sur le disque de centre $(0, 0)$ et de rayon 1 (en vert sur le dessin).

  Soit $z \in \mathbb{R}$:

  $$z = f(x, y) \iff \begin{cases} x^2 + y^2 + z^2 = 1 \\ z \geqslant 0 \end{cases}$$

  Γ_f est la demi-sphère de centre $(0, 0, 0)$, de rayon 1 avec $z \geqslant 0$.



2 Dérivées partielles, gradient

► Pour une fonction de deux variables, on définit deux dérivées : l'une par rapport à la première variable et l'autre par rapport à la deuxième.
On considère la fonction $f : (x, y) \mapsto f(x, y)$, et les applications partielles associées :

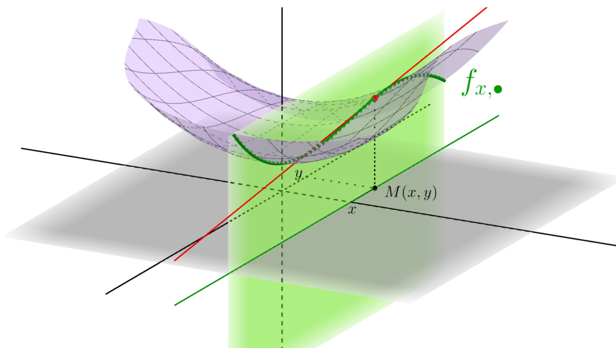
- $f_{\bullet,y} : x \mapsto f(x, y)$, y étant fixé ;
- $f_{x,\bullet} : y \mapsto f(x, y)$, x étant fixé.

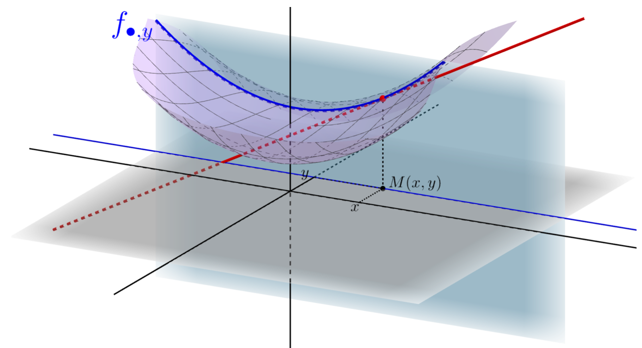
Lorsque ces fonctions sont dérivables, les **dérivées partielles** de la fonction f sont définies de la manière suivante :

$$\boxed{\frac{\partial f}{\partial x}(x, y) = (f_{\bullet,y})'(x) \qquad \text{et} \qquad \frac{\partial f}{\partial y}(x, y) = (f_{x,\bullet})'(y)}$$

ou autrement dit :

$$\frac{\partial f}{\partial x}(x, y) = \frac{\mathrm{d}}{\mathrm{d}x}\big(f(x, y)\big) \qquad \text{et} \qquad \frac{\partial f}{\partial y}(x, y) = \frac{\mathrm{d}}{\mathrm{d}y}\big(f(x, y)\big).$$





⚠ Dans les notations $\frac{\partial f}{\partial x}(x,y)$ et $\frac{\partial f}{\partial y}(x,y)$, le x et le y des fonctions $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ n'ont rien à voir avec le x et le y de leur argument (x,y). Ils signifient que l'on dérive f selon respectivement sa première et sa deuxième variable. On peut par exemple parfaitement parler, si $t \in \mathbb{R}$ de :

$$\frac{\partial f}{\partial x}(2t^2, e^t) \qquad \text{et} \qquad \frac{\partial f}{\partial y}(1-t, t^3)$$

Pour plus de rigueur, les fonctions $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sont parfois notées $\partial_1 f$ et $\partial_2 f$, ou aussi $f^{(1,0)}$ et $f^{(0,1)}$.

Exemples :

- $f : (x,y) \mapsto xy + x^2 - y^3$. On a $\forall x,y,t \in \mathbb{R}$:

$$\frac{\partial f}{\partial x}(x,y) = y + 2x \quad , \quad \frac{\partial f}{\partial y}(x,y) = x - 3y^2 \quad \text{et} \quad \frac{\partial f}{\partial y}(2t^2, e^t) = 2t^2 - 3e^{2t}.$$

- $f : (x,y) \mapsto \sin(xy^2)$. On a $\forall t \in \mathbb{R}$:

$$\frac{\partial f}{\partial x}(3t, e^t) = \Big(\underbrace{(x,y) \mapsto y^2\cos(xy^2)}_{\frac{\partial f}{\partial x}}\Big)(3t, e^t) = e^{2t}\cos(3te^{2t})$$

  et $\forall x,y \in \mathbb{R}$:

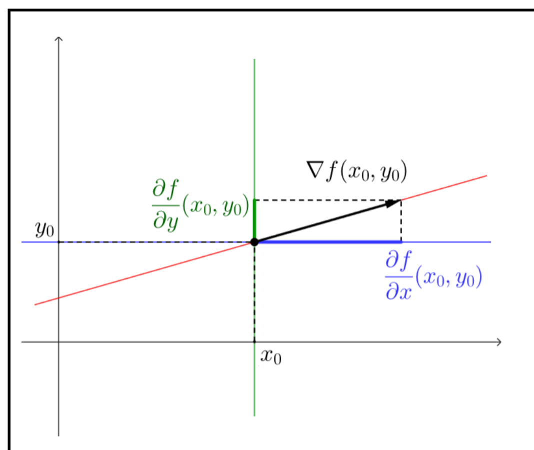
$$\frac{\partial f}{\partial y}(e^{xy}, x^2) = \Big(\underbrace{(x,y) \mapsto 2yx\cos(xy^2)}_{\frac{\partial f}{\partial y}}\Big)(e^{xy}, x^2) = 2x^2e^{xy}\cos\left(e^{xy}x^4\right)$$

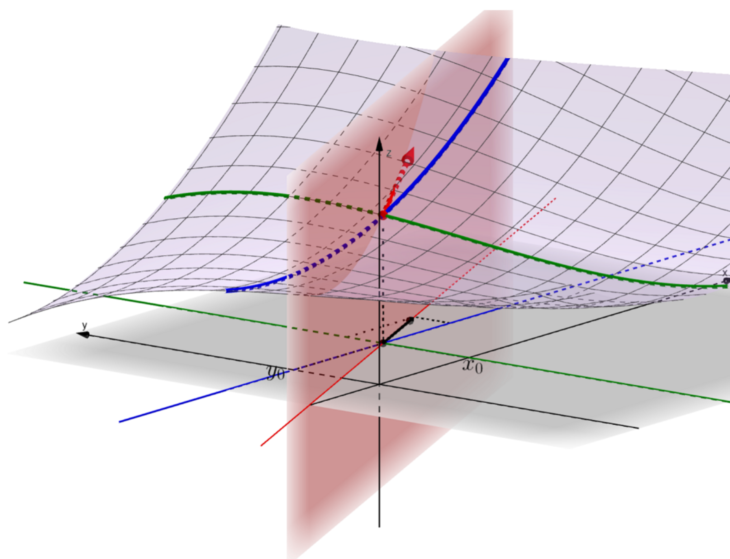
Cf exo 1,2

► Soit $\mathcal{D}'_f$ le sous-ensemble de $\mathbb{R}^2$ sur lequel les fonctions $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sont bien définies. On appelle **gradient** de f, qu'on note ∇f (se lit "gradient de f" ou "Nabla f") l'application définie sur $\mathcal{D}'_f$ par :

$$\boxed{\forall (x,y) \in \mathcal{D}'_f, \quad \nabla f(x,y) = \begin{pmatrix} \frac{\partial f}{\partial x}(x,y) \\ \frac{\partial f}{\partial y}(x,y) \end{pmatrix}}$$

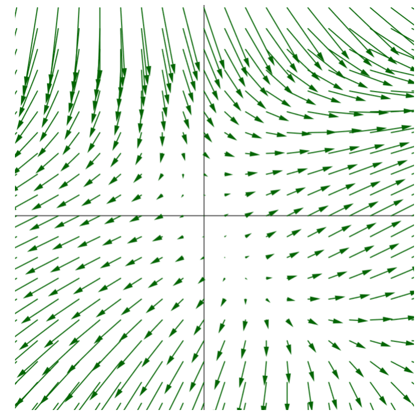
Remarque. *Si l'on fixe un point $M_0 = (x_0, y_0)$ de $\mathcal{D}_f$, le gradient $\boldsymbol{\nabla f(M_0)}$ est un vecteur, que l'on représente souvent en prenant comme origine le point M_0. Il pointe dans la direction où la fonction augmente le plus. En outre, sa norme est égale à la pente de la fonction f dans cette direction précise. On reviendra de façon rigoureuse sur tout cela plus tard dans ce cours.*





Exemples :

- $f : (x,y) \mapsto xy + x^2 - y^3$. $\forall (x,y) \in \mathbb{R}^2$:

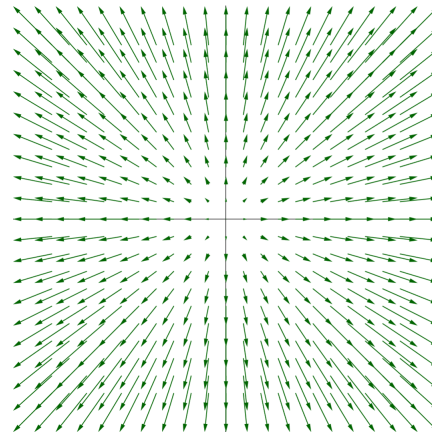


$$\nabla f(x,y) = \begin{pmatrix} y + 2x \\ x - 3y^2 \end{pmatrix}$$

En particulier on a $\nabla f(0,0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\nabla f(1,1) = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$, etc...

Quand M varie, on a un gradient en chaque point. On obtient un "champ de vecteurs".

- $f : (x,y) \mapsto x^2 + y^2$.



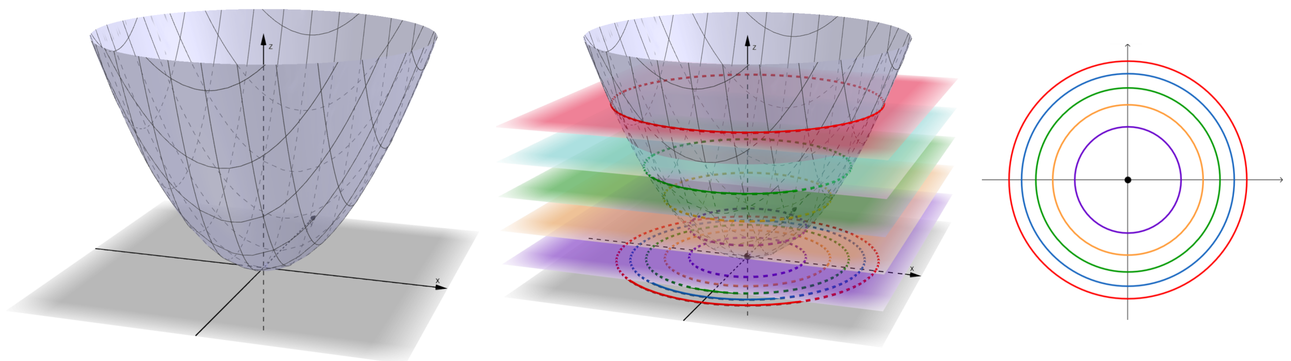
On a pour tout $(x,y) \in \mathbb{R}^2$:

$$\nabla f(x,y) = \begin{pmatrix} 2x \\ 2y \end{pmatrix}$$

3 Lignes de niveau, points critiques (approche graphique)

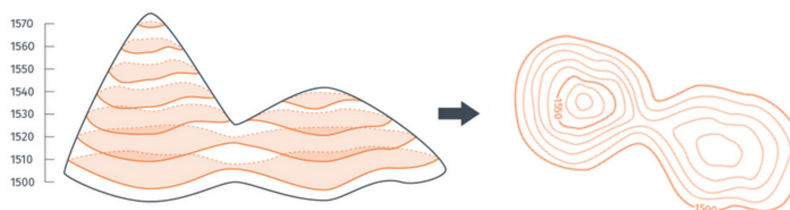
► On appelle **ligne de niveau** $\lambda \in \mathbb{R}$ d'une fonction $f : (x,y) \mapsto f(x,y) \in \mathbb{R}$ la partie de $\mathcal{D}_f$ d'équation $f(x,y) = \lambda$.

Exemple : on considère la fonction $f : (x,y) \mapsto x^2 + y^2$.
Pour $\lambda \in \mathbb{R}_+$, $f(x,y) = \lambda$ est l'équation d'un cercle de centre $(0,0)$ et de rayon $\sqrt{\lambda}$.

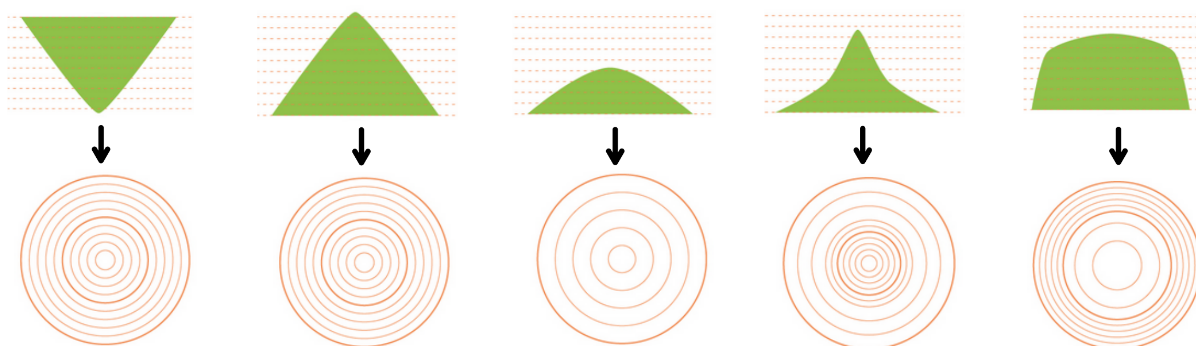


Géométriquement, la ligne de niveau λ s'obtient en considérant l'intersection entre le graphe de f et le plan d'équation $z = \lambda$, puis en projetant (orthogonalement) cette intersection sur le plan d'équation $z = 0$ (c'est-à-dire sur l'ensemble de départ).

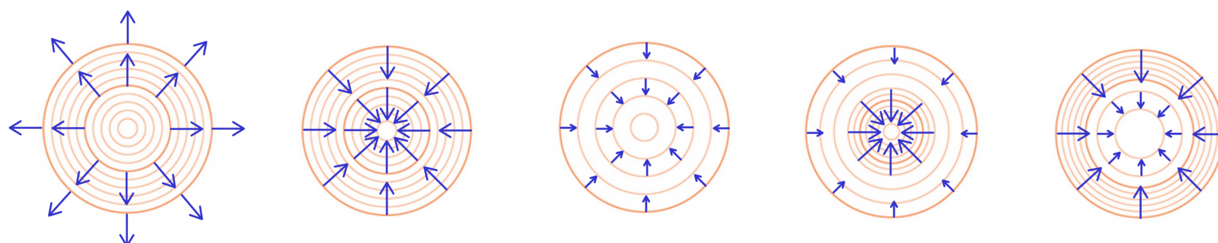
Les lignes de niveau à altitude constante sont représentées sur les cartes topographiques.
Elles permettent d'apprécier le dénivelé sur une carte en 2 dimensions.



Plus les lignes de niveau sont rapprochées, plus la pente est forte, et inversement.

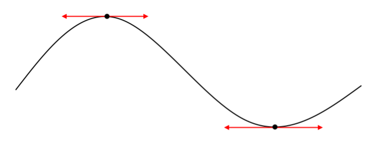


Remarque. *Comme nous le justifierons plus tard, le gradient est orthogonal aux lignes de niveau.*

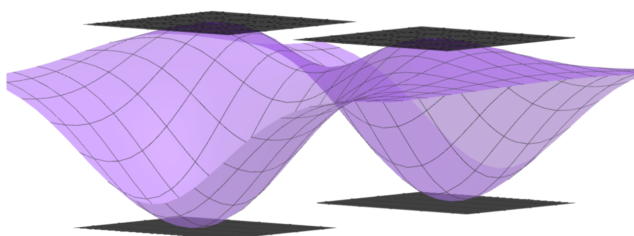


Cf exo 3

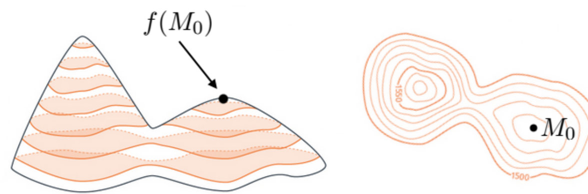
- Quand on cherche les points d'extremum d'une fonction d'une variable qui ne sont pas sur les bords de l'intervalle de recherche, on commence par chercher les points où la dérivée est nulle, et en ces points son graphe admet des droites tangentes horizontales.



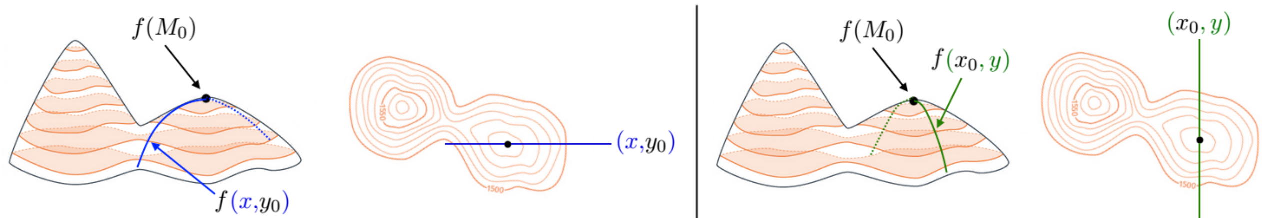
Pour une fonction à deux variables, les points d'extremum (qui ne sont pas aux bords du domaine de recherche) sont des points en lesquels le graphe admet un **plan tangent horizontal**.



Pour les trouver, on cherche les points où **le gradient est nul**. En effet si un point $M(x_0, y_0)$ correspond à un extremum local



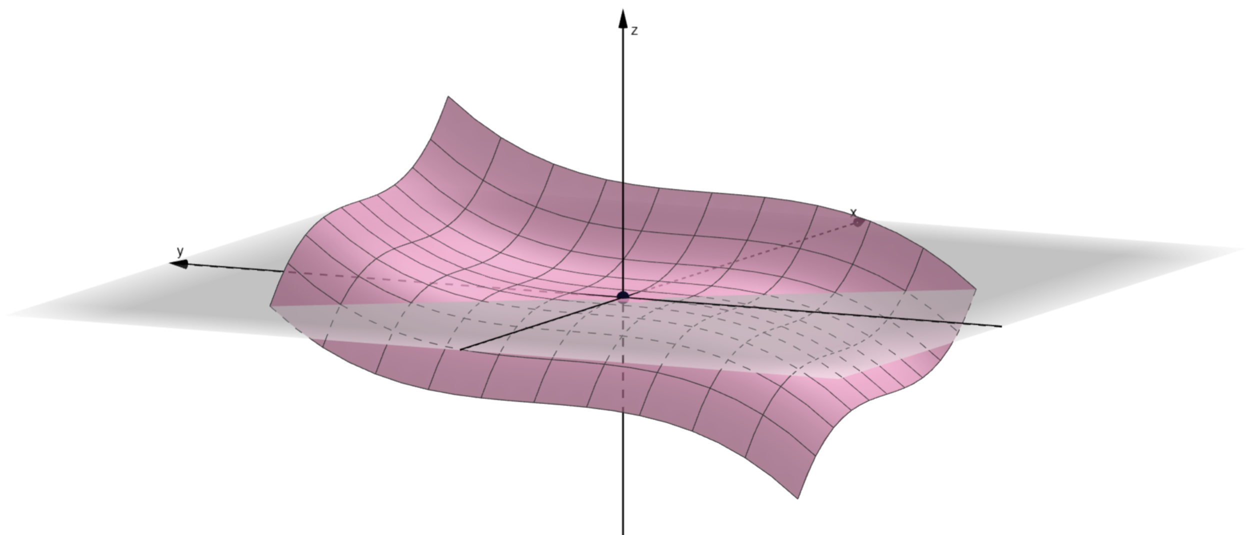
alors les applications partielles $x \mapsto f(x, y_0)$ et $y \mapsto f(x_0, y)$ admettent toutes les deux un extremum local en respectivement $x = x_0$ pour la première, et $y = y_0$ pour la seconde



et donc leurs dérivées s'annulent en ces points ! Autrement dit, $\dfrac{\partial f}{\partial x}(x_0, y_0) = \dfrac{\partial f}{\partial y}(x_0, y_0) = 0$, c'est-à-dire $\boldsymbol{\nabla f(M_0)} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}$.

Ces points sont appelés **points critiques** de la fonction. Comme pour les fonctions d'une variable, **les points d'extremum intérieurs (pas aux bords) d'une fonction sont des points critiques mais ATTENTION la réciproque est fausse**.
Et comme pour les fonctions d'une variable, la recherche des points critiques ne permet pas forcément de trouver tous les extrema car on ne prend pas en compte les bords.
Voici deux contre exemples :

Exemple :
Considérons la fonction $f : (x, y) \mapsto x^3 + y^3$ définie sur $[-1, 1] \times [-1, 1]$.
On a $\forall (x, y) \in [-1, 1]^2, \quad \nabla f(x, y) = \begin{pmatrix} 3x^2 \\ 3y^2 \end{pmatrix}$.
Cherchons les points critiques :
Soit $(x, y) \in [-1, 1]^2, \quad \nabla f(x, y) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow (x, y) = (0, 0)$.
Le point $(0, 0)$ est donc un point critique mais ce n'est pas un point d'extremum :



Par ailleurs, f admet en $(-1, -1)$ son minimum et en $(1, 1)$ son maximum et pourtant ces deux points ne sont pas des points critiques. Ils sont simplement au bord.

Cf exo 4

4 Une approche rigoureuse des notions locales dans $\mathbb{R}^2$

Les notions de continuité, de dérivées partielles et de gradient d'une fonction f en un point $X_0 = (x_0, y_0)$ de $\mathbb{R}^2$ ne dépendent pas seulement de la valeur de f en X_0 mais aussi de son comportement au *voisinage* de X_0. Pour aborder ces questions de façon rigoureuse, c'est cette notion de *voisinage* qu'il faut éclaircir. En dimension 2 (et en fait en dimension quelconque), elle se traduit par la notion de *boule* autour de X_0.

On se place dans $\mathbb{R}^2$ muni de la norme euclidienne canonique : pour $X = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$, $\|X\| = \sqrt{x^2+y^2}$.

4.1 Ouverts

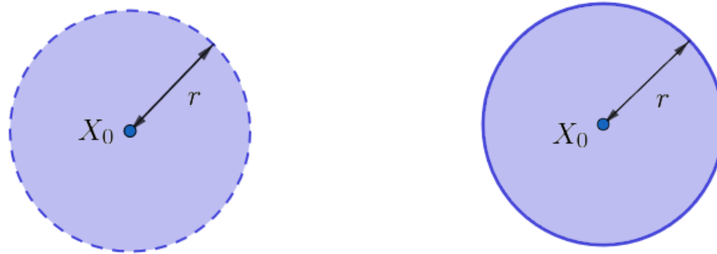
Définition (Boule ouverte, boule fermée)**.**
Soit $X_0 = (x_0, y_0)$ un point de $\mathbb{R}^2$ et r un réel strictement positif.

- *On appelle* ***boule fermée*** *de centre X_0 et de rayon r, et on note $B_f(X_0, r)$, l'ensemble :*

$$\boldsymbol{B_f(X_0, r) = \{X = (x, y) \in \mathbb{R}^2 \ \ \textbf{tq} \ \ \|X - X_0\| \leqslant r\}}.$$

- *On appelle* ***boule ouverte*** *de centre X_0 et de rayon r, et on note $B(X_0, r)$, l'ensemble :*

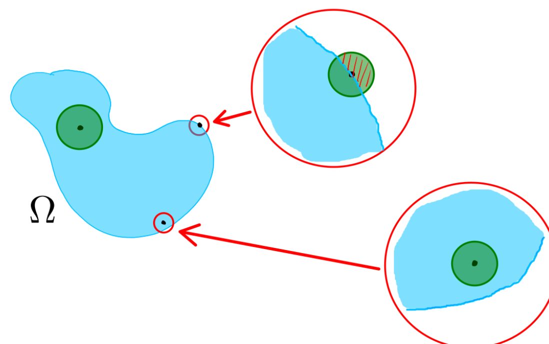
$$\boldsymbol{B(X_0, r) = \{X = (x, y) \in \mathbb{R}^2 \ \ \textbf{tq} \ \ \|X - X_0\| < r\}}.$$



Remarque. *Les boules ouvertes et les boules fermées de $\mathbb{R}^2$ généralisent les intervalles ouverts $]a, b[$ et les intervalles fermés $[a, b]$ de $\mathbb{R}$.*

Définition (Ouverts et fermés de $\mathbb{R}^2$)**.**
Soit Ω une partie de $\mathbb{R}^2$.

- *On dit que Ω est un* ***ouvert*** *de $\mathbb{R}^2$ lorsque, pour tout point X_0 de Ω, il existe une boule ouverte de centre X_0 entièrement incluse dans Ω.*
  Autrement dit : $\forall X_0 \in \Omega, \quad \exists r > 0 \quad tq \quad B(X_0, r) \subset \Omega$.
- *On dit que Ω est un* ***fermé*** *de $\mathbb{R}^2$ lorsque son complémentaire est un ouvert.*



Remarques.

- *Intuitivement, une partie de $\mathbb{R}^2$ est ouverte lorsqu'elle ne contient aucun "point frontière", et elle est fermée lorsqu'elle contient tous ses "points frontières".*
- *Tout comme il existe des parties de $\mathbb{R}$ qui ne sont ni ouvertes ni fermées (par exemple, les intervalles du type $]a,b]$), il existe des parties de $\mathbb{R}^2$ qui ne sont ni ouvertes ni fermées.*
- *Considérer une fonction f définie sur un ouvert de $\mathbb{R}^2$ est une situation confortable d'un point de vue théorique, car chaque point X_0 possède autour de lui un voisinage, c'est-à-dire une boule. On va donc pouvoir introduire des notions locales, comme la limite, la continuité, et les dérivées partielles.*

Dans la suite du cours, Ω désigne un ouvert de $\mathbb{R}^2$.

4.2 Continuité

Définition (Limite)**.**
Soit $f : \Omega \to \mathbb{R}$ une fonction et $X_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \in \Omega$.
On dit que $f(X)$ tend vers $\ell \in \mathbb{R}$ lorsque $X = \begin{pmatrix} x \\ y \end{pmatrix}$ tend vers X_0 et on note $f(X) \xrightarrow[X \to X_0]{} \ell$ si

$$\forall \varepsilon > 0,\ \exists \eta > 0 \ \ tq \ \ \forall X \in B_f(X_0, \eta),\ |f(X) - \ell| \leqslant \varepsilon$$

c'est-à-dire si

$$\boxed{\forall \varepsilon > 0,\ \exists \eta > 0 \ \ tq \ \ \forall X \in \Omega,\ \ \|X - X_0\| \leqslant \eta \Rightarrow |f(X) - \ell| \leqslant \varepsilon}.$$

Définition (Continuité)**.**
Soit $f : \Omega \to \mathbb{R}$ une fonction.

▶ *On dit que f est continue en $X_0 \in \Omega$ si $f(X) \xrightarrow[X \to X_0]{} f(X_0)$.*

▶ *On dit que f est continue sur Ω si elle est continue en tout point de Ω.*

Remarque. *Cette définition étant finalement la même que celle pour les fonctions d'une variable, on obtient (et avec les mêmes démonstrations) les théorèmes d'opérations sur les fonctions continues :*

- *la somme, le produit, le quotient (quand le dénominateur est non nul) de fonctions continues sont continus ;*
- *la composée de fonctions continues est continue.*

4.3 Fonctions de classe $\mathcal{C}^1$

Définition (Fonction de classe $\mathcal{C}^1$)**.**
Si f admet des dérivées partielles en tout point de Ω et que les fonctions $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ sont continues sur Ω, alors on dit que f est ***de classe*** *$\mathcal{C}^1$ sur Ω.*
On note $\mathcal{C}^1(\Omega, \mathbb{R})$ ou $\mathcal{C}^1(\Omega)$ l'ensemble des fonctions de classes $\mathcal{C}^1$ définies sur Ω.

Remarque. *De la même manière que précédemment :*

- *une fonction de classe $\mathcal{C}^1$ sur Ω est continue sur Ω.*
- *la somme, le produit, le quotient (quand le dénominateur est non nul), la composée de fonctions de classe $\mathcal{C}^1$ sont de classe $\mathcal{C}^1$.*

4.4 Extrema

Définition (extremum)**.**
Soit D une partie non vide de $\mathbb{R}^2$, $f : D \to \mathbb{R}$ une fonction et $X_0 \in D$. On dit que :

▶ *f admet un* ***maximum*** *(respectivement* ***minimum****) en X_0 si :*
$\forall X \in D, \quad f(X) \leqslant f(X_0)$ *(respectivement $f(X) \geqslant f(X_0)$).*

▶ *f admet un* ***maximum local*** *(respectivement* ***minimum local****) en X_0 si :*
$\exists \eta > 0 \ \ tq \ \ \forall X \in D \cap B(X_0, \eta), \quad f(X) \leqslant f(X_0)$ *(respectivement $f(X) \geqslant f(X_0)$).*

▶ *f admet un* ***extremum*** *(respectivement local) en X_0 si f admet un maximum ou un minimum (respectivement local) en X_0.*

Théorème (extremum et point critique)**.**
Soit $f \in \mathcal{C}^1(\Omega, \mathbb{R})$ et $X_0 \in \Omega$ (Ω est un ouvert, donc X_0 est donc un point intérieur).
***Si** f admet un extremum local en X_0, **alors** X_0 est un point critique de f c'est-à-dire $\boldsymbol{\nabla f(X_0)} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}$.*

Démonstration. Supposons que X_0 est un point d'extremum local de f, mettons un maximum local. Considérons un réel $r > 0$ suffisamment petit pour que :
- $B(X_0, r) \subset \Omega$ (possible puisque Ω est un ouvert) ;
- pour tout $X \in B(X_0, r)$, $f(X) \leqslant f(X_0)$ (possible puisque X_0 est un point de maximum local).

Si un réel x est suffisamment proche de x_0 (typiquement $x \in]x_0 - r, x_0 + r[$), alors le point (x, y_0) appartient à la boule $B(X_0, r)$, et donc $f(x, y_0) \leqslant f(x_0, y_0)$.
Cela montre que l'application partielle $x \mapsto f(x, y_0)$ admet en $x = x_0$ un maximum local. Par théorème (sur les fonctions à une variable), sa dérivée s'annule donc en x_0. Donc $\frac{\partial f}{\partial x}(x_0, y_0) = 0$.
Symétriquement, on a $\frac{\partial f}{\partial y}(x_0, y_0) = 0$, et donc finalement : $\nabla f(x_0, y_0) = 0_{\mathbb{R}^2}$. □

Remarque. *Nous avons vu précédemment que la réciproque est fausse.*

5 Développement limité à l'ordre 1 et plan tangent

Théorème (Développement limité à l'ordre 1)**.**
*Soit $f \in \mathcal{C}^1(\Omega, \mathbb{R})$ et $X_0 \in \Omega$. Alors, f admet un développement limité d'ordre **1** en X_0.*
On a : (pour H appartenant à un voisinage de $0_{\mathbb{R}^2}$ suffisamment petit pour que $X_0 + H \in \Omega$) :

$$\boldsymbol{f(X_0 + H) \underset{H \to 0_{\mathbb{R}^2}}{=} f(X_0) + \nabla f(X_0) \cdot H + \underbrace{o(\|H\|)}_{=\|H\| o(1)}}$$

où $o(1)$ est une fonction de deux variables qui tend vers 0 lorsque H tend vers 0. Le "·" entre $\nabla f(X_0)$ et H désigne le produit scalaire usuel.

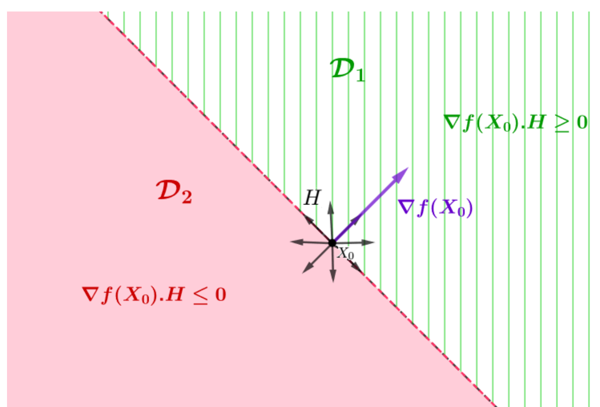
Remarques.
- *La traduction en physique de ce théorème est :*

$$df = \overrightarrow{grad}(f) \cdot \overrightarrow{dr}$$

En effet :

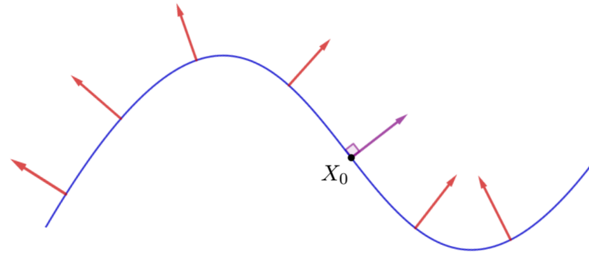
$$\underbrace{f(X_0 + H) - f(X_0)}_{df} = \underbrace{\nabla f(X_0)}_{\overrightarrow{grad}(f)} \cdot \underbrace{H}_{\overrightarrow{dr}} + \underbrace{o(\|H\|)}_{\text{absent à l'ordre 1}}$$

- *Il signifie qu'à l'ordre 1, la variation de f entre X_0 et $X_0 + H$ vaut $\boldsymbol{\nabla f\,(X_0) \cdot H}$.*



Cette variation est donc :
- ▶ *maximale quand H est colinéaire de même sens que $\nabla f(X_0)$;*
- ▶ *minimale quand H est colinéaire à $\nabla f(X_0)$ de sens contraire ;*
- ▶ *positive sur $\mathcal{D}_1$;*
- ▶ *négative sur $\mathcal{D}_2$;*
- ▶ *nulle sur la droite passant par X_0 et orthogonale à $\nabla f(X_0)$.*

Ce raisonnement local explique pourquoi globalement les lignes de niveau sont orthogonales aux gradients de la fonction. Se déplacer orthogonalement aux gradients laisse f constant.



- *Autre formulation du théorème : en notant $X_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ et $H = \begin{pmatrix} h \\ k \end{pmatrix}$, on a :*

$$\boldsymbol{f(x_0 + h, y_0 + k) \underset{\begin{pmatrix} h \\ k \end{pmatrix} \to \begin{pmatrix} 0 \\ 0 \end{pmatrix}}{=} f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0) \times h + \frac{\partial f}{\partial y}(x_0, y_0) \times k + o\left(\left\| \begin{pmatrix} h \\ k \end{pmatrix} \right\|\right)}$$

ou encore, avec le changement de variables $h = x - x_0$ et $k = y - y_0$:

$$f(x, y) \underset{\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}}{=} f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0) \times (x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0) \times (y - y_0) + o\left(\left\| \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} \right\|\right)$$

- *Cette dernière formulation permet d'affirmer que la fonction affine :*

$$(x, y) \mapsto f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0) \times (x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0) \times (y - y_0)$$

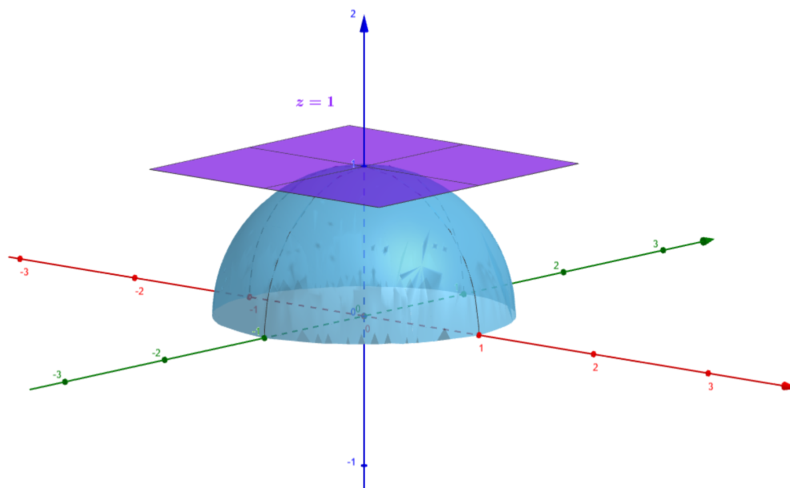
approche au premier ordre la fonction f au voisinage du point (x_0, y_0).
En particulier, le plan d'équation

$$z = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0) \times (x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0) \times (y - y_0)$$

qui est le graphe de cette fonction affine, est le ***plan tangent*** *du graphe de la fonction f en (x_0, y_0). (C'est-à-dire à la surface d'équation $z = f(x, y)$). On retrouve ici une formule analogue à celle de l'équation de la tangente en x_0 pour les fonctions d'une variable :*

$$y = f(x_0) + f'(x_0)(x - x_0)$$

Exemple : Reprenons l'exemple de la demi-sphère et déterminons l'équation du plan tangent au sommet. On a $f(0,0) = 1$ et $\nabla f(0,0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ donc $\boldsymbol{z = 1}$ est l'équation du plan tangent à la demi-sphère au point $(0, 0)$.



Cf exo 5

6 Règle de la chaîne

Théorème (Règle de la chaîne)**.**
Soit $f \in \mathcal{C}^1(\Omega, \mathbb{R})$*,* I *un intervalle de* $\mathbb{R}$ *et* $(\gamma_1, \gamma_2) \in \left(\mathcal{C}^1(I, \mathbb{R})\right)^2$*.*
On suppose que $\gamma : t \mapsto (\gamma_1(t), \gamma_2(t))$ *est à valeurs dans* Ω*.*
Alors $f \circ \gamma : t \mapsto f(\gamma(t)) = f(\gamma_1(t), \gamma_2(t))$ *est de classe* $\mathcal{C}^1$ *et :*

$$\forall t \in I, \qquad \frac{d}{dt}(f(\gamma(t)) = (f \circ \gamma)'(t) = \nabla f(\gamma(t)) \cdot \gamma'(t)$$

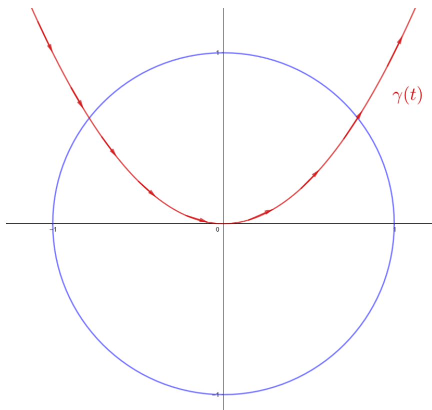
où $\gamma'(t) = (\gamma_1'(t), \gamma_2'(t))$*.*

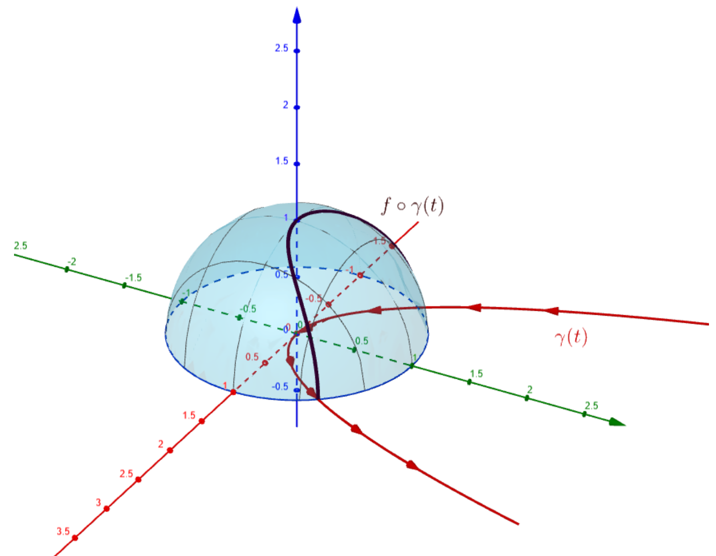
Démonstration. (Ou plutôt idée de la démonstration) On fixe $t_0 \in I$, et on compose le DL_1 de γ en t_0 avec le DL_1 de f en $X_0 = \gamma(t_0)$ pour obtenir un le DL_1 de $f \circ \gamma$ en t_0. Cela montre que $f \circ \gamma$ est dérivable en t_0 et nous donne la valeur de $(f \circ \gamma)'(t_0)$. □

Remarque. *Autre formulation, plutôt pour les calculs pratiques :*

$$\frac{d}{dt}(f(\gamma_1(t), \gamma_2(t)) = \frac{\partial f}{\partial x}(\gamma_1(t), \gamma_2(t)) \times \gamma_1'(t) + \frac{\partial f}{\partial y}(\gamma_1(t), \gamma_2(t)) \times \gamma_2'(t)$$

Exemple : On considère $\gamma : t \mapsto (t, t^2)$ et $f : (x, y) \mapsto \sqrt{1 - x^2 - y^2}$ définie sur le disque de centre $(0, 0)$ et de rayon 1.





$\gamma'(t) = (1, 2t)$ et d'après la règle de la chaîne :

$$\begin{aligned}
(f \circ \gamma)'(t) &= \nabla f(\gamma(t)) \cdot \gamma'(t) \\
&= \frac{-\gamma_1(t)}{\sqrt{1 - \gamma_1(t)^2 - \gamma_2(t)^2}} \times \gamma_1'(t) + \frac{-\gamma_2(t)}{\sqrt{1 - \gamma_1(t)^2 - \gamma_2(t)^2}} \times \gamma_2'(t) \\
&= \frac{-t(1 + 2t^2)}{\sqrt{1 - t^2 - t^4}}.
\end{aligned}$$

Cf exo 6

Remarque. *Si au lieu de donner en paramètres à f deux quantités dépendant chacune d'une seule variable, $\gamma_1(t)$ et $\gamma_2(t)$, on donne deux quantités dépendant chacune de deux variables, c'est-à-dire si l'on regarde une fonction du type :*

$$g \ : \ (u,v) \mapsto f\left(\varphi(u,v), \psi(u,v)\right)$$

alors les dérivées partielles et le gradient de g se déduisent de la formule de la chaîne.
En effet, si l'on s'intéresse à $\dfrac{\partial g}{\partial u}$, alors il s'agit de dériver $f\left(\varphi(u,v), \psi(u,v)\right)$ par rapport à u, en considérant dans l'opération que v est fixée. Ainsi, dans cette opération, les quantités $\varphi(u,v)$ et $\psi(u,v)$ sont des quantités qui ne dépendent que d'une seule variable : u. La règle de la chaîne s'applique. On obtient alors :

$$\begin{aligned}\frac{\partial g}{\partial u}(u,v) &= \frac{d}{du}\left[f\left(\varphi(u,v), \psi(u,v)\right)\right] \\ &= \frac{\partial f}{\partial x}\left(\varphi(u,v), \psi(u,v)\right) \times \frac{\partial \varphi}{\partial u}(u,v) + \frac{\partial f}{\partial y}\left(\varphi(u,v), \psi(u,v)\right) \times \frac{\partial \psi}{\partial u}(u,v).\end{aligned}$$

De même, si l'on dérive par rapport à v :

$$\begin{aligned}\frac{\partial g}{\partial v}(u,v) &= \frac{d}{dv}\left[f\left(\varphi(u,v), \psi(u,v)\right)\right] \\ &= \frac{\partial f}{\partial x}\left(\varphi(u,v), \psi(u,v)\right) \times \frac{\partial \varphi}{\partial v}(u,v) + \frac{\partial f}{\partial y}\left(\varphi(u,v), \psi(u,v)\right) \times \frac{\partial \psi}{\partial v}(u,v).\end{aligned}$$

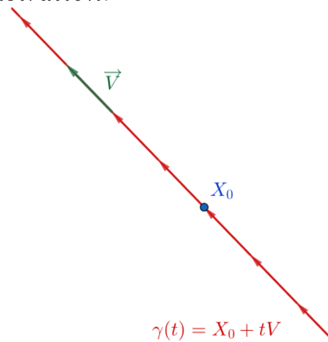
Cf exo 7

Proposition (Dérivée selon un vecteur)**.**
Soit $f \in \mathcal{C}^1(\Omega, \mathbb{R})$, $X_0 \in \Omega$ et $V = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^2$.
L'application $g_V \ : \ t \mapsto f(X_0 + tV)$ est définie au voisinage de 0, dérivable en 0 et de dérivée :

$$\boldsymbol{g'_V(0) = \nabla f\left(X_0\right) \cdot V = \frac{\partial f}{\partial x}(x_0, y_0) \times v_1 + \frac{\partial f}{\partial y}(x_0, y_0) \times v_2.}$$

On l'appelle ***dérivée de f en X_0 selon le vecteur V.***

Démonstration.



C'est un cas particulier de la règle de la chaîne.
$\gamma : t \mapsto X_0 + tV = (x_0 + tv_1, y_0 + tv_2)$ est bien de classe $\mathcal{C}^1$.
$\gamma'(t) = (v_1, v_2)$ et $(f \circ \gamma)'(0) = \nabla f(X_0) \cdot V$.

□